图片内容元数据体系研究与标准编制

摘　要：图片已经成为连接媒体、品牌以及消费者的最高效的传播媒介之一。图片特征描述是图片检索、推荐和交换的重要基础。基于当前图片应用和图片元数据标准的现状分析，文章提出了着重描述图片内容和语义特征的图片内容元数据体系框架，并以此为基础完成了国家标准《中文新闻图片内容描述元数据规范》的编制，填补了国内图片元数据标准的空白，有助于提高图片标注效率和质量，更好地满足图片获取和交易交换的需要。

关键词：元数据；图片内容元数据；国家标准
中图分类号：G203　　文献标识码：A
文章编号：1671-0134（2017）11-081-04　　DOI：10.19483/j.cnki.11-4653/n.2017.11.026

文 / 张　弛　黄　菁　张鹏洲　武国卫

图片在新闻、出版、广告、营销等各领域扮演着重要角色。尤其在融合媒体不断发展的背景下，图片凭借其信息量大、视觉吸引力强、数据量小等优势，在两微一端应用、精品内容生产等方面发挥越来越大的作用。发现和获取高质量图片是图片应用的核心环节，图片特征表示是支撑这一核心环节的基石。图像视觉特征自动提取技术不断取得进步，近年来深度学习技术在手写体字符识别、人脸识别、以图搜图、图像分割等图像处理领域取得了突破性进展[1]。另一方面，基于文本的图片元数据仍然是图片特征描述的主要手段，随着文本处理和自然语言处理技术的不断发展，其研究与应用具有广阔的发展空间。

《中国互联网络发展状况统计报告》指出，移动互联网主导地位不断强化，以即时通信、搜索引擎、网络新闻为代表的基础互联网应用用户规模趋于稳定，提供精准优质内容服务为重点方向[2]。图片成为提供精准优质内容不可或缺的最高效的传输媒介，微信朋友圈每日上传图片超过 10 亿张，视觉中国图片库拥有编辑类图片超过 1.4 亿张，每天新增超过 20000 张，个人、商业和媒体机构等拥有的各类图片库规模也在不断增长。

图片需要被发现和流转才能发挥价值，这要求图片要能被较好地描述和存储。针对图片元数据应用现状，笔者对国内多家大规模商业图片库和媒体机构图库进行了实地调研。调研发现：（1）随着图片数量增长和流动速度加快，图片检索的难度也越来越大；（2）现有图片元数据标准不能很好满足图片标引和检索的需要，尤其缺乏与内容语义相关的元数据；（3）图片标引现状较为混乱，元数据标准不统一，自定义的私有元数据情况非常普遍、标引过程缺乏质量控制机制等；（4）对统一的图片元数据规范的需求迫切。

由此，本文提出了一种面向图片主题和内容语义描述的图片内容元数据体系，加强标注内容语义完整性，提高标注规范性和标注结果的兼容性和一致性，并以此为基础，完成了国家标准《中文新闻图片内容描述元数据规范》的编制。

1. 研究基础

对图片应用、构图和编辑的研究，从不同角度加深对图片的理解，是构建图片内容元数据体系的重要基础。

1.1 图片应用

目前主流的图片库大致分为以下几种类型。一是大型综合商业图片库，如视觉中国、东方 IC 等，拥有大量专业签约摄影师，主要为机构用户提供高质量的编辑类图片、创意类图片和微视频资源等。二是媒体机构图片库，如中国全球图片总汇以及各大报社图片库等，以编辑类图片为主，拥有各自的优势资源。三是提供各类设计素材的图库，如全景网、壹图网、昵图网、懒人图库、拍信等。四是聚焦垂直领域，如电商、汽车、地理、摄影、网页素材等。此外还有以个人分享交流为重心的社交类图片库等。无论是哪一类图片库，重中之重是充当好连接图片创作者和消费者之间的桥梁。而发现图片是这架桥梁的价值体现。

通过对代表性图片库系统的对比分析，图片检索功能主要可分为以下几种：

关键词检索：应用普遍，一般与图片主题、画面事物、人物、地点、事件的相关度较高。

按类别检索：类别划分区别较大。除了常规意义的知识领域分类法外，基于标签的浅层化的图片主题分类被越来越多地采用。

专题事件检索：广泛应用于新闻图片组织和查询。

图片属性筛选：提高图片检索准确率，有助于快速定位图片。常用属性包括：图片来源（个人 / 机构）、是否授权、

色彩、尺寸、片幅 / 画幅、人物信息（人数、年龄、性别、种族）、画面背景等。

图片推荐：主要有热门推荐、编辑推荐、相似推荐、个性化推荐等。

精细化检索：主要实现方式有：在结果中检索；对上次检索结果集合中图片所标注的标签进行统计倒排，供用户点选；相关标签推荐。

通过对图片检索功能的分析总结发现：

（1）使用文字来查找图片仍然是图片检索的最主要方式。关键词、图片类别、专题事件、检索过滤用的图片属性，甚至图片推荐，都离不开文字标注信息。

（2）图片特征描述的维度数量越来越多。强化图片特征描述，增加维度、提高精度（例如由简单的图说细化为主题词列表等），弥补传统图片编目属性数量少的不足，以更好地适应从海量图片中快速定位到目标图片的客观需求。

（3）现有图片元数据标准未较好满足实际应用需要，普遍存在自定义私有图片元数据项的情况。有些是共性需求，有些是个性化需要。

（4）图片内容和语义特征的价值凸显。画面内容、表现形式和主题概念，是图片的灵魂。几乎所有图片检索功能都与图片内容有关。

1. 2 图片构图

构图是指将现实生活中的物体利用视觉特征，再现到二维空间内，通过画面构造传达阐释更多的信息，反映作者对事物的认识和感情，并具有突出主体、吸引视线、简化杂乱、使画面和谐均衡等作用。构图强调如何把人、景、物等安排在画面中以获得最佳布局，在形式美方面诉诸于视觉的点、线、形态、用光、色彩等的配合。构图的目的是传达信息、表达主题以及创作者的认知和感情。

从构图的角度来看，一幅画面主要包括主体、陪体和环境三部分。主体是画面的主要表现对象，往往是画面的结构中心和视觉中心。陪体是指在与主体构成一定的情节，帮助表达主体的特征和内涵的对象。环境分为前景和背景，对主体和情节起烘托作用。处理好主体、陪体和环境的关系是表达主题和图片质量评价的关键。此外，用光、影调、摄影技巧等在图片创作中不可或缺。

1. 3 图片编辑

1.3.1 报道摄影

《图片编辑手册》[3]中明确界定了报道摄影（Photojournalism），即媒体上图片与文字相结合的报道形式，它涵盖媒体中出现的多种形式的摄影报道。创意摄影、特写新闻报道属于报道摄影的不同表现形式，报道摄影有别于宣传摄影（公关摄影）和画意摄影，形象性和新闻性是报道摄影的两大特性。

优秀的报道摄影应具备历史价值、社会价值、心理作用和美学价值。一张照片带给读者的不仅是视觉内容本身，还包含通过视觉内容所表现出的主题、视觉美感或冲击、情感、意境等隐含信息。

1.3.2 照片评价

新闻照片早期的评判标准是“新、真、活、情、意”。随着传媒的不断发展，对于报道摄影图片的评价标准也演变为技术标准、信息传递、美学标准、传播效果。技术标准（包括曝光、色温、景深、焦点等）为基础，关注照片传递语义信息是否准确丰富，更关注传播效果的好坏（从内容到形式都对传播效果有重要影响）。

1.3.3 照片说明写作

图片说明要对图片中容易被察觉的事物和不容易察觉但较重要和有趣的细节予以说明，同时避免倾向性和主观评价。单幅照片说明大多用两句话完成。第一句讲画面上正在发生什么，交代时间、地点、人物和简单事件描述；第二句讲相关背景。组图说明写作方法包括直接说明事件信息、从事件多个侧面中选择、突出事件背景介绍等。一般而言，单图说明侧重点明人物、场景和正在发生什么，组图说明侧重背景介绍和深入细致描写。

由此可见，不论是报道摄影特征、照片评价标准，还是照片说明，关注的都不仅仅是画面自身的视觉内容，所以从技术、美学、事件、背景、重要细节等多维度来实现多方位、立体式的图片特征描述具有重要意义。

2. 图片内容元数据体系

2. 1 图片元数据层次划分

元数据是关于数据的数据。在数字图书馆领域，元数据被分为描述型、管理型和结构型三类。《IPTC Photo Metadata》[4]作为权威的图片元数据标准在国际上得到广泛采用，它将图片元数据划分为描述（Descriptive）和技术（Technical）两类。本文基于语义抽象程度将图片元数据划分为物理层、逻辑层、内容层和主题层，如图 1 所示。

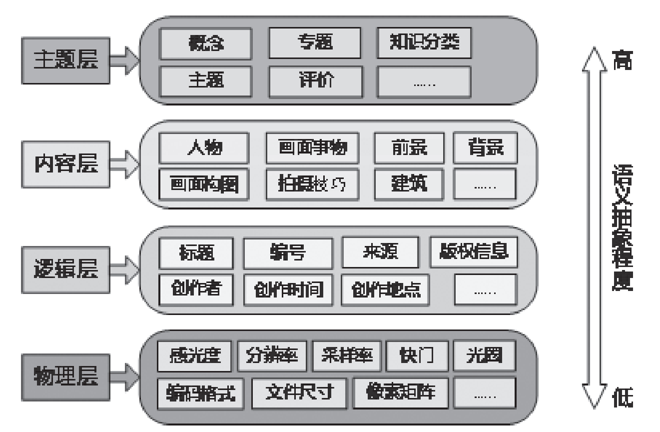


图 1　图片元数据层级划分

物理层元数据主要包括数字图片文件的相关属性、图片拍摄时的技术参数信息和图片底层视觉特征。逻辑层元数据包括图片自身应用属性和授权信息等。内容层和主题层在物理层和逻辑层之上，主要描述图片所呈现出的画面内容和主题信息，其中内容层侧重画面所见事物，主题层侧重于通过画面内容所表达出的概念和思想。

对上述四层进一步归类，将物理层和逻辑层统称为属性元数据，内容层和主题层统称为内容元数据。本文针对后者提出了图片内容元数据体系。

2. 2 图片内容元数据体系

《IPTC Photo Metadata》聚焦图片描述元数据，主要涉及逻辑层、内容层元数据，先后发布了多个版本，元数据间缺乏结构关系，并且未涉及主题层元数据。EXIF 图片标准主要侧重图片物理层元数据。《中文新闻信息置标语言》[5] 主要面向新闻稿件，对图片特征描述未作细化的规约。再结合前文的分析可见，构建面向图片内容和语义的图片元数据体系和编制标准规范在图片应用、图片特性、现实情况等多方面都是需要的，具有重要应用价值。

如何界定图片内容？看到一张图片后，读者直观感受到的画面中的事物、色彩、构图等属于图片内容范畴，同时这些视觉信息使读者感受、联想、体会到的意境、思想、情感等也是图片内容的重要组成部分。

如何描述图片内容？以本文“研究基础”部分所述为基础，采用自底向上和自顶向下相结合的思路，最终形成三位一体的体系结构。

图片内容元数据包括形、技、意三个方面。“形”是视觉表现，“技”是创作手法，“意”是图片主题和灵魂。“形”和“技”为“意”服务，“意”通过“形”和“技”来表达，三者不是相互孤立的，而是相辅相成、相互融合而成为一个有机整体。

具体而言，“形”是指一幅画面中的主要事物及其特征、环境信息等。“技”是指照片拍摄技术有关的信息。“意”是指图片主题、语义以及图片所属知识领域分类等信息。三个维度包含若干个子维度，如图 2 所示。因大多维度可以顾名思义，所以在此不对每个维度的具体含义作展开说明。

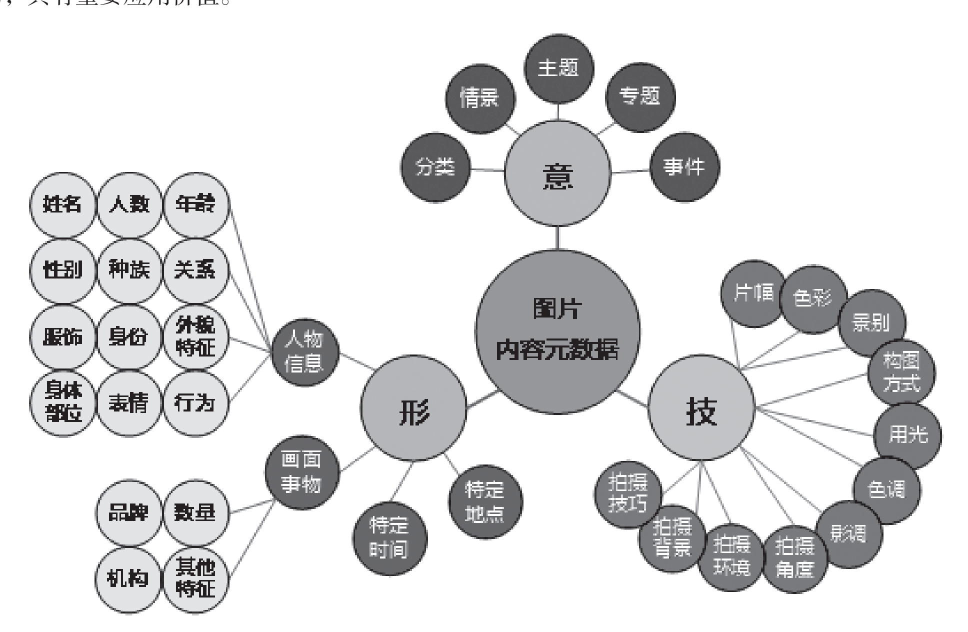


图 2 图片内容元数据体系

3. 标准编制

《中文新闻图片内容描述元数据规范》于 2014 年 9 月获得国家标准委批准立项，笔者作为项目组成员，主要负责标准文档的起草编制工作。标准编制得到了来自标准研制机构、媒体机构、商业图片库、学术组织、相关科技公司等多位标准领域和图片领域专家的大力支持和指导。本标准在图片内容元数据体系的基础上，经过反复调研、撰写、评审、反馈、修改，目前已完成报批稿上报国家标准委。

3. 1 编制原则

（1）适用性。以满足业务需求为出发点，突出内容语义特征描述，减少层级，提高标注时采用受控词表的比例。

（2）协调性。对与已有相关标准中含义相同或相近的元数据元素采用了一致的命名，避免在多个标准应用过程中的混淆。

（3）专业性。面向图片内容和语义特征描述，元数据元素主要围绕视觉内容、主题概念、拍摄技术三个角度，有多位领域专家的具体指导。

3. 2 标准内容

《中文新闻图片内容描述元数据规范》从新闻编辑用图和创意配图两个应用角度定义了中文新闻图片内容和语义特

征描述的元数据元素集合，适用于图片数据的采集、编辑、存储、发布、检索、交换等处理环节。

元数据元素集合分为三个部分：公共元数据、新闻编辑类图片元数据、新闻创意类图片元数据，共包括 40 个元数据元素。其中“人物信息”和“摄影技术”包含二级元素。如图 3 所示。

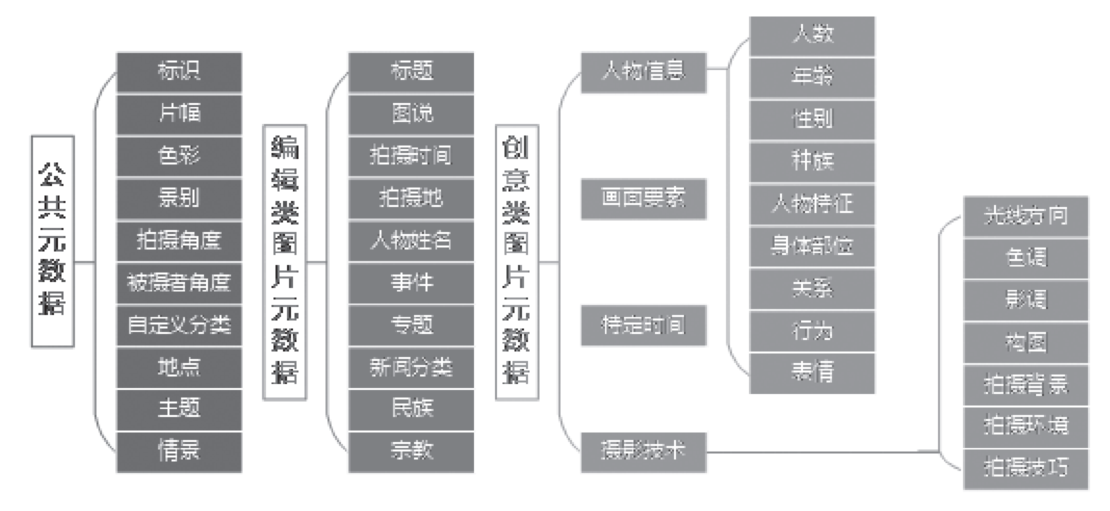


图 3 《中文新闻图片内容描述元数据规范》元数据集合

对比图 2 和图 3 可发现，标准中增加几个逻辑层元数据元素，如标题、图说、拍摄时间、拍摄地和标识。实际应用中这几项是必选的，增加的目的是为了应用方便。公共元数据部分的元数据元素是对应到所有类别图片的，所以将片幅、色彩、景别等从“摄影技术”元素中取出纳入到公共元数据部分。

标准中规定的必填项为 8 个，集中在公共和编辑类图片元数据部分。标准中给出了 22 个受控词表，用于 20 个元数据元素标注。元数据的详细定义说明等内容详见标准正文。

3.3 标准应用建议说明

3.3.1 标注时注重规范性

基于受控词表有助于提高标注的准确性和一致性，标准支持受控词表扩展。对于自由文本标注的元数据元素，在实际标注时也建议自定义标注规范。

3.3.2 标注时注重完整性

标准中规定为必填项的元数据元素较少，更好地满足要求图片快速发布的需要。建议采用多级标注方式，粗粒度和细粒度标注相结合。此外，应用中可细化对某些元数据元素的标注要求，例如“画面主体”，对于时尚类图片的“品牌”和动物图片的“品种”等可增设二级元数据。

3.3.3 自动标注与人工标注相结合

利用计算机图像处理技术，实现如“色彩”“人数”“色调”“影调”等的自动标注。从图片描述文字中自动提取关键词，有助于提高对“主题”“地点”“时间”等维度的标注效率。

标准编制的基本目标是更好地满足图片检索和交换的需要。图片多维度标注的结果集合也是图片内容分析、资源关联和聚合、相似图片推荐等的重要基础。

结语

图片内容元数据体系从形、技、意三个维度建立内容和语义特征描述框架，以此为基础编制了《中文新闻图片内容描述元数据规范》，符合实际应用需要和图片自身特性的要求。标准中未涉及图片版权的元数据定义，原因在于图片版权与图片内容相关度不高，建议在图片应用中直接参照现有版权相关标准执行。

图片内容特征描述在图片应用中具有重要作用。以深度学习为代表的计算机视觉处理技术不断取得突破性进步，如何基于文本的图片特征描述与自动处理技术深度结合以更好地满足多场景下的图片业务需求是具有重要研究和应用价值的课题。媒

参考文献

[1] 刘建伟，刘源，罗雄麟 . 深度学习研究进展 . 计算机应用研究 . 第 31 卷第 7 期 . 2014.7.

[2] 中国互联网络信息中心 . 第 40 次中国互联网络发展状况统计报告 . 2017（7）.

[3] 任悦，曾璜编著 . 图片编辑手册（第四版）. 中国摄影出版社 . 2015（9）.

[4]IPTC Photo Metadata Standard. http：//www.iptc.org/std/photometadata/specification /IPTC-PhotoMetadata.

[5] 全国中文新闻信息标准化技术委员会 . 国家标准 GB/T 2009-2013《中文新闻信息置标语言》.

（作者单位：新华社通信技术局）